2016

A Comparative Analysis of an Interior-point Method and a Sequential Quadratic Programming Method for the Markowitz Portfolio Management Problem

Zhifu Xiao
Oberlin College

A Comparative Analysis of an Interior-point Method and a Sequential Quadratic Programming Method for the Markowitz Portfolio Management Problem

Zhifu Xiao

March 30, 2016

Abstract

In this paper, I give a brief introduction of the general optimization problem as well as the convex optimization problem. The portfolio selection problem, as a typical type of convex optimization problem, can be easily solved in polynomial time. However, when the number of available stocks in the portfolio becomes large, there might be a significant difference in the running time of different polynomial-time solving methods. In this paper, I perform a comparative analysis of two different solving methods and discuss the characteristics and differences.

1 Introduction

In this paper, I discuss the convex optimization problems and an important property of convex optimization problems. By solving Markowitz portfolio management problem using two different methods, I compare two commonly-used algorithms on this topic. The major contribution is a comparative analysis of two widely-used methods for solving nonlinear convex optimization problems.

This paper is organized as follows. In section 1.1 I introduce the general optimization problem and the convex optimization problem. I discuss the important feature of convex optimization and prove why is it important. In section 1.2 I talked about some applications of convex optimization problems. In section 1.3 I present a brief literature review of the background of Markowitz portfolio management problem and some related works. In section 1.4 I give a small example to explain the model and the results of this model. In section 2 I mainly explain the data and methods I used to solve Markowitz portfolio management problem. In section 3 I report on my implementation of the models and my comparative analysis. In the last section I conclude the results I have and dress some further research interests.

1.1 Convex optimizations: Theory

Optimization, or specifically, mathematical optimization, is mathematical approaches to maximizing or minimizing a function by systematically choosing input values within the designated sets and computing the function values [1]. Mathematically, we can represent the general optimization problem as:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq b_i, \ i = 1, \ldots, m \end{array} \tag{1}$$

Here, the vector of $x = (x_1, \ldots, x_n)$ is the list of the *optimization variables* of the problem. In this general form of optimization problems, the function $f_0 : \mathbf{R}^n \to \mathbf{R}$ is the *objective function*. The functions $f_i(x), i = 1, \ldots, m$, are *constraint functions*, and the constraints $b_1, \ldots, b_m$, are limits or bounds for the constraints. We define a vector x^* to be the *optimal solution* of the problem (1) if x^* has the smallest objective value among all vectors satisfying the constraints: for any x with $f_i(x) \leq b_i,\ i = 1, \ldots, m,\ f_i(x) \geq f_i(x^*)$ holds for any x.

A *convex optimization problem* is an optimization problem in which the objective function and constraint functions are convex. A function f is *convex* if and only if for any $x_1, x_2 \in \mathbf{R}^n$ and any scalar λ with $0 \leq \lambda \leq 1$, we have

$$f[\lambda x_1 + (1-\lambda)x_2] \leq \lambda f(x_1) + (1-\lambda)f(x_2)].$$

Another important notion is that of a *convex set*. A set C is *convex* if and only if for any $x_1, x_2 \in C$ and any scalar θ with $0 \leq \theta \leq 1$, we have

$$\theta x_1 + (1-\theta)x_2 \in C.$$

Convexity is very important for optimization problems, but what exactly are the reasons why convexity is crucial? First, why is having a convex objective function important? A very fundamental and important property of the unconstrained convex optimization problem is that any local optimal solution is also the global optimal solution. The proof is shown below.

Proof. (by contradiction) Suppose x^* is a local optimal point for an unconstrained convex optimization problem, we know x^* is definitely feasible and

$$f_0(x^*) = \inf\{f_0(\widehat{x}) \mid \widehat{x} \text{ feasible},\ \|\widehat{x} - x^*\|_2 \leq R\} \tag{2}$$

for some $R > 0$. Suppose x^* is not a global optimal solution, then there exist a feasible point $\overline{x}$ such that $f_0(\overline{x}) < f_0(x^*)$. Obviously, we have $\|\overline{x} - x^*\|_2 > R$ since if otherwise, $f_0(x^*) \leq f_0(\overline{x})$. Create a point $\widehat{x}$ such that

$$\widehat{x} = (1-\lambda)x^* + \lambda\overline{x},\quad \lambda = \frac{R}{2\|\overline{x} - x^*\|_2}.$$

For $\widehat{x}$, we have $\|\widehat{x} - x^*\|_2 = R/2 < R$. By the definition of convexity, $\widehat{x}$ is feasible and we have

$$f_0(\widehat{x}) \leq (1-\lambda)f_0(x^*) + \lambda f_0(\overline{x}) < f_0(x^*).$$

This result contradicts equation (2) because it violated the assumption that x^* is a local optimal point. □

Thus, we prove this important property that for an unconstrained convex optimization problem, any local optimal solution is also a global optimal solution. However, if we have a constrained convex optimization problem, we need both the objective function and all of the constraint functions to be convex in order to guarantee this important property. Why is convexity of the constraint functions important? If only the objective function is convex, what could happen?

Convexity of the constraint functions is important because if the constraint functions are convex, then the feasible region, which is the set of points that satisfy all constraints, will be a convex set.

Proof. Suppose for the optimization problem, we know the constraint functions are convex and we have $f_i(x) \leq b_i,\ i = 1, \ldots, m$. We firstly want to show that $C = \{x : f_i(x) \leq b_i\}$ is a convex set for $i = 1, \ldots, m$.

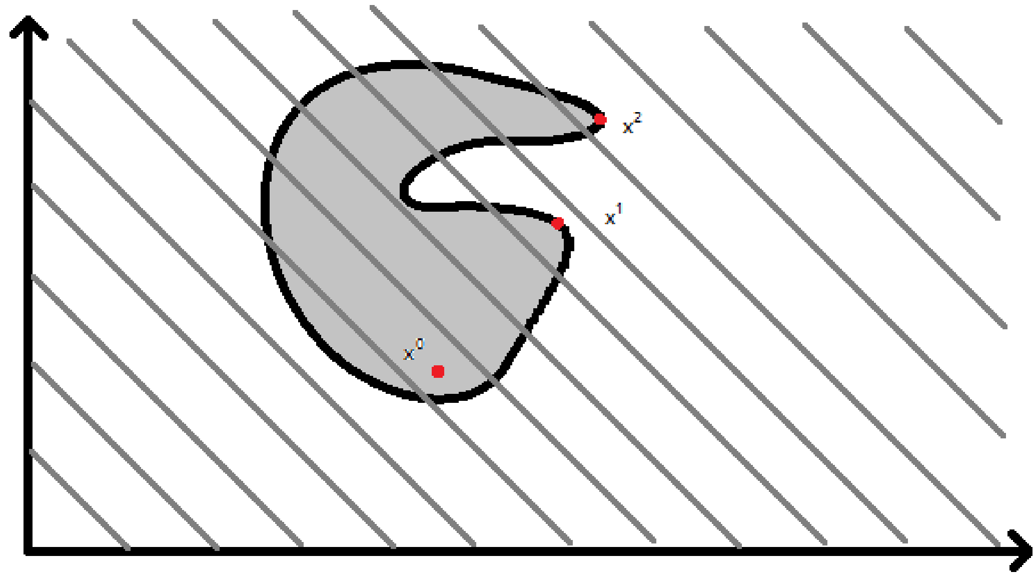
To prove $\{x : f_i(x) \leq b_i\}$ is a convex set for $i = 1, \ldots, m$, we want to show that for any $u, v \in C$, $\theta\, u + (1-\theta)\, v \in C_i$ for every $0 \leq \theta \leq 1$. Let $\overline{x}$, $\widehat{x} \in C$ and $0 \leq \theta \leq 1$. We have

$$\begin{aligned} f(\theta\,\overline{x} + (1-\theta)\,\widehat{x}) &\leq f(\theta\,\overline{x}) + f((1-\theta)\,\widehat{x}) \\ &= \theta\, f(\overline{x}) + (1-\theta)\, f(\widehat{x})) \\ &\leq \theta\, b_i + (1-\theta)\, b_i \\ &= b_i \end{aligned}$$

Thus, $f(\theta\overline{x} + (1-\theta)\widehat{x}) \leq b_i$, which means that the set $C = \{x : f_i(x) \leq b_i\}$ is a convex set for $i = 1, \ldots, m$.

Second, we want to show that the intersection of convex sets is also a convex set. Let set C_i, $i = 1, \ldots, m$ be the convex sets corresponding to the constraints. Let $\mathcal{C} = \cap C_i$, $i = 1, \ldots, m$. Then, for any $x_1, x_2 \in \mathcal{C}$, $x_1, x_2 \in C_i$, $i = 1, \ldots, m$. Since we know that C_i are convex sets, by definition, for any θ with $0 \leq \theta \leq 1$, $\theta x_1 + (1-\theta)x_2 \in C_i$. Hence, $\theta x_1 + (1-\theta)x_2 \in \mathcal{C}$. □

In conclusion, convexity is important for both objective function and constraint functions: the convexity in objective function ensures that if the feasible region is a convex set, any local optimal solution is also the global optimal solution; the convexity in constraint functions ensures that the feasible region is a convex set and neither is dispensable. On the other hand, if only the objective function is convex but constraint functions are not convex, we will end up with a non-convex feasible region and we cannot obtain the important property that the local optimal is also global optimal.



In the graph, the feasible region is shaded and several contours of the objective function are shown. If the initial point starts at x^0, it will move along the edge and finally arrive at x^1. It will get stuck at x^1, the local optimal because the objective function value will be decreasing in all directions. However, the global optimal is actually in x^2.

1.2 Convex optimizations: Applications

Convex optimization problems are used in a variety of environments: In statistics, the ordinary least square (OLS) regression is a basic type of unconstrained convex optimization problem. Suppose we have a dataset consisting of n paired observations (x_i, y_i), $i = 1, \ldots, n$, where x_i is a $m \times 1$ vector of m variables that we want to use to predict y_i, which is the response variable, based on several observations. An ordinary approach is to minimize the squared errors between the real value of y_i and our linear predictions $\widehat{y_i} = A_i x_i + b_i$, where A_i is a $1 \times m$ vector of coefficients on all variables x_i and b_i is a constant term. The general form of an OLS regression is shown below:

$$\text{minimize} \quad \sum_{i=1}^{n} \|\widehat{y}_i - y_i\|_2^2 \tag{3}$$

To prove that (3) is a convex function, we need to prove that the sum of convex functions is also a convex function. The proof is shown below.

Proof. Suppose we have $f(x)$ and $g(x)$ and both of them are convex function. We want to show that $z(x) = f(x) + g(x)$ is also a convex function. Let $\overline{x}$, $\widehat{x} \in R^n$ and $0 \le \theta \le 1$. We have

$$\begin{aligned} z(\theta\,\overline{x} + (1-\theta)\,\widehat{x}) &= f(\theta\,\overline{x} + (1-\theta)\,\widehat{x}) + g(\theta\,\overline{x} + (1-\theta)\,\widehat{x}) \\ &\le f(\theta\,\overline{x}) + f((1-\theta)\,\widehat{x})) + g(\theta\,\overline{x}) + g((1-\theta)\,\widehat{x}) \\ &= f(\theta\,\overline{x}) + g(\theta\,\overline{x}) + f((1-\theta)\,\widehat{x})) + g((1-\theta)\,\widehat{x}) \\ &= z(\theta\,\overline{x}) + z((1-\theta)\,\widehat{x}) \end{aligned}$$

We showed that $z(\theta\,\overline{x} + (1-\theta)\,\widehat{x}) \le z(\theta\,\overline{x}) + z((1-\theta)\,\widehat{x}))$. Thus, $z(x)$, the summation of two convex function, is also a convex function. □

As a result, we could easily see that the objective function (3) is a convex function and the OLS regression is an unconstrained convex optimization problem. This OLS regression could be easily solved by an analytical solution since it is a particularly simple unconstrained convex optimization problem. Compared to OLS regression, Markowitz portfolio optimization is the type of constrained convex optimization problem that cannot be solved in analytical way. However, it is still a convex optimization problem and we could solve it in polynomial time by using some algorithms.

1.3 Portfolio selection: background and previous work

Before Harry Markowitz's article discussing the the two stages in the process of selecting a portfolio, there were several approaches to portfolio selection to maximize the expected return. Harry Markowitz first wrote about his theory of portfolio selection in 1952, and he later published a book in 1959. In his book and article, he made two important assumptions: first, investor does (or should) maximize discounted expected returns; second, the investor does (or should) consider expected return a desirable thing and variance of return as an undesirable thing [2]. Based on these two assumptions, he rejected the previous hypothesis that investor does (or should) maximize discounted return. The trade-off between mean return and variance is a fundamental theme that appears over and over again in mathematical approaches to investment.

After Markowitz's work, Sharpe and Lintner introduced the capital asset pricing model (CAPM) for the valuation of assets, a model that describes the relationship between expected return and risk. They introduced an important measurement of risk: Beta. It is now widely used to compare a certain investment to the market average: If beta equals to 0, we will consider this certain investment to be a "risk-free" investment. If beta is less than 1, it means that the risk of this investment is less than the market average. Otherwise, this investment is considered to be more risky than the market average [3, 4].

In the 1970s and 1980s, scholars worked to put Markowitz's theory in the real-world. B. Blog designed a small portfolio selection model for small investor so that they could calculate their optimal portfolios in a short amount of time [5]. Jobson and Korkie published a paper to improve the reliability of Markowitz's theory by using some relaxations over the implementation to calculate the portfolio quickly [6]. Other approaches to Markowitz's theory includes solving the problem by using a more efficient polynomial-time algorithm [7], proving the sufficient conditions for the convergence of the solving algorithms [8], adding fixed transaction costs into the model and solve it [9] and designing a algorithm for large-scale portfolio optimization [10].

In recent years, accompanied with the rapid growth of computation ability, lots of scholars have started to use more advanced methods that requires larger storage and complicate computations. H. Tanaka introduced fuzzy probabilities and possibility distributions in Markowitz's model to reduce the variance in the optimal portfolio [11]. O. Ledoit reduced the variance through a different approach: he used weighted average of two covariance matrix to estimate the covariance matrix of stock returns and reduce the covariance of optimal portfolios [12]. In 2007, L. Garlappi used Bayesian methods to create a model with multiple priors and aversion to ambiguity. The optimal portfolios from his model are more stable and yield a higher Sharpe ratio compared to classical and Bayesian models [13]. At the same time, Jun Liu analysed this problem by using explicit portfolio weights to solve dynamic portfolio choice problems and re-balance the portfolios [14]. C. Harvey also used Bayesian decision theoretic framework to create higher order moments in portfolio selection that lead to higher expected return [15].

1.4 Portfolio selection: Models and an example

In this section, I discuss the basic Markowitz's portfolio selection model and introduce some variations. Then I illustrate a small example based on the model and solve it using the basic convex optimization method.

Suppose there are n assets available to select for the portfolio, let x_i represents the amount of asset i held in the time period and denote the price change of asset i during that time period as P_i. For now I only allow long positions in the model, which means that x_i is non-negative. For each asset i, we consider P_i to be a random variable with mean $\overline{P}_i$. I define M to be the covariance matrix such that $M_{ij} = \mathrm{cov}(P_i, P_j)$. I could now define this problem and the approach as a trade-off between expected return and variance: I want to minimize the variance of the portfolio while requiring that the portfolio gain a minimum expected return. The optimization problem can be represented as the following:

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} x_i M_{ij} x_j \\ \text{subject to} & \displaystyle\sum_{i=1}^{n} \bar{P}_i x_i \geq r_{\min} \\ & \displaystyle\sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0, i = 1, \ldots, n \end{array}$$

$r_{\min}$ denotes the expected return I want the portfolio to reach. Here the variables are x_i; the data are the expected value $\overline{P}_i$ and the covariance matrix M. It can be easily proved that this optimization problem is a convex quadratic optimization problem since the objective function and all constraint functions are all convex quadratic functions. I can use vector form to represent this problem in a clear manner as the following:

$$\begin{array}{ll} \text{minimize} & x^T M x \\ \text{subject to} & \text{-}\,\bar{P}^T x \leq \text{-}\,r_{\min} \\ & e^T x = 1,\ x \succeq 0 \end{array}$$

From now on, I will use the matrix form. Let's look at a tiny example with the number of available stocks in the portfolio, n, equals to 10. Then the covariance matrix M is a 10×10 matrix, the expected return $\bar{P}$ is a 10×1 vector and the variable x is also a 10×1 vector. Before talking about the method I used to solve this problem, first let's look at the results to develop a sense for what the model does. Here is a table representing the value of the objective function with respect to different settings of the minimum expected return $r_{\min}$.

Asset	Expected Return $\bar{P}$	Optimal Sol. ($r_{\min} = 0.05$)	Optimal Sol. ($r_{\min} = 0.15$)	Optimal Sol. ($r_{\min} = 0.25$)
1	0.1485	-	0.0983	-
2	0.1350	0.1476	0.7856	0.0263
3	0.0489	0.2671	-	-
4	0.0550	-	-	-
5	0.0289	0.5756	-	-
6	0.2531	-	0.1161	0.9737
7	0.1105	-	-	-
8	0.0434	-	-	-
9	0.0432	-	-	-
10	0.0382	0.0097	-	-
Objective Function Value		0.08635	0.40837	3.8786

Table 1: Results for different settings of $r_{\min}$

By looking at this table, it is shown that the stock with the highest expected return among all 10 stocks is stock 6 with $\bar{P}_6 = 0.2531$ and the stock with the smallest expected return among all 10 stocks is stock 5 with $\bar{P}_5 = 0.0289$. From the result, it illustrates that when $r_{\min}$ is relatively small, the portfolio is well-balanced by selecting moderate numbers of stocks. When $r_{\min}$ becomes larger, there is similar needs to balance the portfolio, but there will be more risky stocks to fulfill the need of minimum expected return requirement. When $r_{\min}$ becomes really large, the portfolio will be heavily skewed to the most risky assets. Here is a graph of the efficient frontier of this small portfolio management problem where r_{min} is between 0 and 0.25. Note that this model when $n = 10$ is very tiny and could be easily solved in 1 second. In the next section I will consider a larger portfolio and no longer consider the situation where $n = 10$.

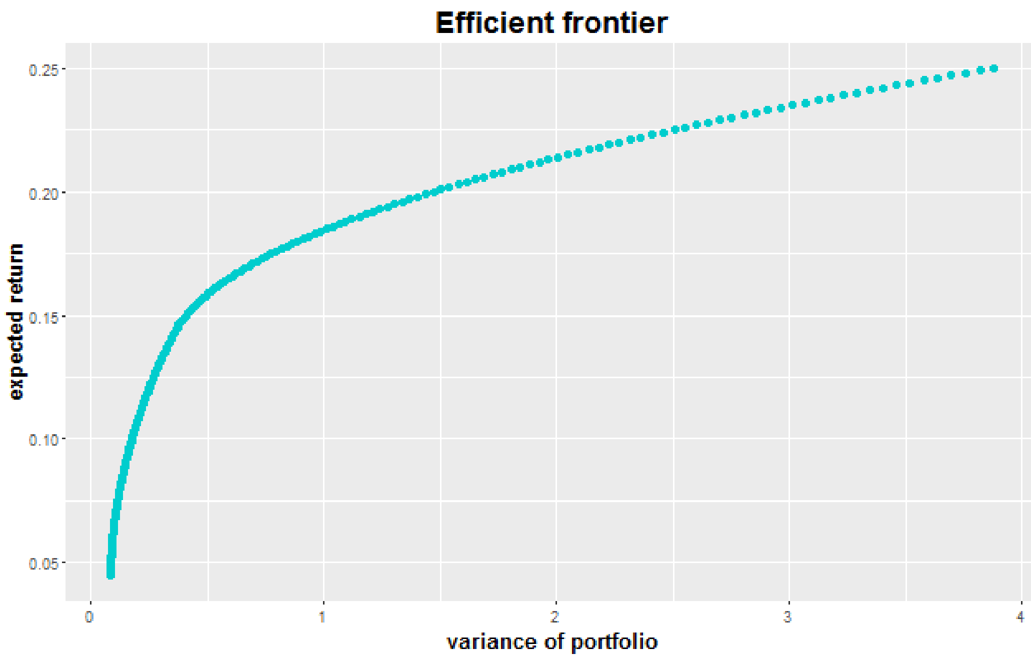


Figure 1: The efficient frontier of this tiny Markowitz problem with $n = 10$

From the graph, it is clear that when r_{min} is less than 0.15, the variance of the portfolio grows slowly by the increasing of r_{min}. When r_{min} is greater than 0.15, the variance of the portfolio increases rapidly, as the portfolio needs to become more aggressive so that it could reach the minimum expected return requirement. In the later section I will give a brief explanation and explain how the algorithm works.

2 Data and methods

In this section, I introduce the dataset I used for the portfolio management optimization problem and the methods I used to clean and analyze the dataset. I briefly talk about the algorithms I used to solve this problem and some characteristics of the algorithms.

2.1 Data

I used open-source Yahoo! Finance API to gather daily market adjusted closed data of thousands of stocks in 130 trading days(half calender year). There are a total of 1411 stocks available in the portfolio and I created the covariance matrix and expected return for those stocks. Here is a brief table representing some basic statistics of my data.

Min.	Q_1	Median	Mean	Q_3	Max.	Std. Dev.
-0.4830	-0.0080	0.0212	0.0265	0.0498	0.6168	0.0631

Table 2: Summary statistics for $\bar{P}$

The table shows that the majority of data lies between -0.008 and 0.0498 with some large outliers. The standard deviation is small, which means that the data is mostly concentrated in the center.

2.2 Methods

There are several available algorithms to solve inequality constrained convex optimization problems and here I discuss two different approaches: interior-point method and sequential quadratic programming (SQP) method. Both methods require the objective function and constraint functions to be twice differentiable. The first and second derivative of the objection function are shown below.

$$\begin{aligned} \nabla f(x) &= \frac{\partial f(x)}{\partial x} \\ &= 2Mx \end{aligned}$$

$$\begin{aligned} \nabla^2 f(x) &= \frac{\partial^2 f(x)}{\partial x^2} \\ &= 2M \end{aligned}$$

Once the first and second derivative of the objective function are available, either interior-point method or SQP method could be implemented to optimize the value of the objective function.

The first polynomial-time interior point linear programming (convex optimization problems in which the objective function and all constraint functions are linear) was devised by Karmarkar in 1984 [7]. It was then developed to adapt to more different scenarios, like several kinds of nonlinear optimization problems. In the 1990s, it expanded to semi-definite optimization and now a days, it is a mature polynomial-time algorithm for nonlinear optimization problems.

Sequential quadratic programming method was firstly invented by Wilson in his PhD thesis in 1963 [16]. It is a solving method for some kinds of nonlinear optimization problems where objection function and all constraints functions are twice continuously differentiable. SQP method is the generalization of Newton's method, which is a method for unconstrained optimization problems by iteratively searching the roots of twice continuously differentiable objective function. By solving quadratic sub-problems, SQP method could give a convincing result in polynomial time.

Those two methods are widely-used in academia as well as in industry, and comparing those two methods, one of the advantage of SQP method is that the initial point and all iterative steps are not required to be feasible points, while interior-point method need to keep track of the feasible region so that it could bypass the infeasible region. However, a modification of interior-point method, which is the infeasible-interior-point method, allows for problems with no strictly feasible points by increasing the computational complexity.

3 Results and analysis

In this section, I briefly explain the implementation of the Markowitz portfolio management problem when the number of available stocks becomes large. Then I discuss some methods to deal with the initial setup of the portfolio. Finally, I display the results and do the comparative analysis.

3.1 Results

I used R to call the API provided by Yahoo! Finance and downloaded the daily market data as I described before. After downloading the data, I used R to clean the data and match the date. Then I calculated the expected return and covariance matrix over the time period. By using Gurobi and Matlab, I implemented the algorithms to solve the convex quadratic optimization problem and plot the computing time versus the number of available stocks. Here I did a randomization of my portfolios 100 times and created a 95% confidence interval on the running time. I also calculated the confidence interval for the objective function values, but the difference between the two methods is not significant. So I used the mean value of variance instead of the confidence interval of variance. Results and outputs are shown below.

n	$r_{\min} = 0.05$		$r_{\min} = 0.15$		$r_{\min} = 0.25$	
	Variance	Time(s)	Variance	Time(s)	Variance	Time(s)
100	0.0197	(0.082, 0.086)	0.1753	(0.04, 0.063)	0.6054	(0.059, 0.095)
1000	0.0020	(28.433, 32.775)	0.0489	(10.343, 17.161)	0.2489	(6.904, 12.73)

Table 3: Results for different settings of $r_{\min}$ under SQP method

In this table, it is shown that under SQP method, as the number of available stocks increases, the computing time increases significantly. Also, when $n = 100$, the difference of computing time between different $r_{\min}$ is very little; when n increases from 100 to 1000, the difference of computing time between different $r_{\min}$ increases a lot. There exist some overlaps when $n = 1000$: The 95% confidence interval when $r_{\min} = 0.15$ is $(10.343, 17.161)$, while the 95% confidence interval when $r_{\min} = 0.25$ is $(6.904, 12.73)$. It seems that the variance of the computing time becomes larger when n increases.

n	$r_{\min} = 0.05$		$r_{\min} = 0.15$		$r_{\min} = 0.25$	
	Variance	Time(s)	Variance	Time(s)	Variance	Time(s)
100	0.0197	(0.151, 0.159)	0.1754	(0.103, 0.111)	0.6056	(0.089, 0.099)
1000	0.0022	(36.342, 38.317)	0.0493	(34.058, 35.219)	0.2505	(24.021, 25.566)

Table 4: Results for different settings of $r_{\min}$ under interior-point method

Under interior-point method, as the number of available stocks increases, the computing time increases significantly just as SQP method does. However, the difference is that although the mean running time under interior-point method is larger than that under SQP method, the range of the confidence interval under interior-point method is smaller than that under SQP method. There is no such overlap of confidence interval under interior-point method, which means interior-point method is more stable for calculating the optimal solution; SQP method has faster speed but also has larger variance.

From the results in the two tables above, as $r_{\min}$ increases, the computing time tends to decrease in general. Also, as the number of available stocks increases, the computing time increases significantly in both methods. When $r_{\min}$ is relatively small, the portfolio needs to be well-diversified and thus increases the number of steps toward the optimal solution. As $r_{\min}$ increases, the portfolio tends to be more concentrated to high-return and high-variance stocks to fulfill the minimum expected return requirement. Since my initial setup of the portfolio is the stock of nearest expected return to $r_{\min}$, it will take less steps from the initial point to the optimal point.

Note that I set up the initial portfolio by checking the expected return of all available stocks and setting the proportion of the asset that has the closest expected return to r_{min} equals to 1. The advantage of doing so is that when $r_{\min}$ is large, the initial portfolio is close to the optimal portfolio and will take less steps and time to compute the optimal solution. If we randomly setup the portfolio, it will take more time for both algorithms to converge to the optimal solution. Results are shown below and note that here I use the average time rather than the confidence interval of running time.

N	$r_{\min} = 0.05$		$r_{\min} = 0.15$		$r_{\min} = 0.25$	
	Variance	Time(s)	Variance	Time(s)	Variance	Time(s)
100	0.0197	0.1403	0.1753	0.0709	0.6054	0.0952
1000	0.0020	33.3961	0.0489	17.4862	0.2489	12.0742

Table 5: Results of SQP method under the random initial portfolio

N	$r_{\min} = 0.05$		$r_{\min} = 0.15$		$r_{\min} = 0.25$	
	Variance	Time(s)	Variance	Time(s)	Variance	Time(s)
100	0.0197	0.1778	0.1754	0.1117	0.6056	0.0988
1000	0.0022	38.4058	0.0493	35.8012	0.2505	26.2018

Table 6: Results of interior-point method under the random initial portfolio

It is shown that the mean running time under the random initial portfolio is around or larger than the higher end of the 95% interval. As n becomes larger, the running time for randomized portfolio is still longer than the previous portfolio setting. It shows that the selection of initial point is important in both algorithms to speed up the computing time. However, it is possible that there exist some other initialization methods that works better than the previous setting. I discuss it later in the conclusion section and it could be further examined.

3.2 Analysis

After comparing two algorithms in two specific settings ($n = 100$ and $n = 1000$), I compare the computing time versus the number of available stocks between two algorithms under the same $r_{\min}$ and I compare the difference between different value of $r_{\min}$. Note that here I use a specific setting of the portfolio rather than do randomization for 100 times. As the results, there exists some outliers but not significant. Figures are shown below and here n ranges from 10 to 1000 increasing by 10 in each step.

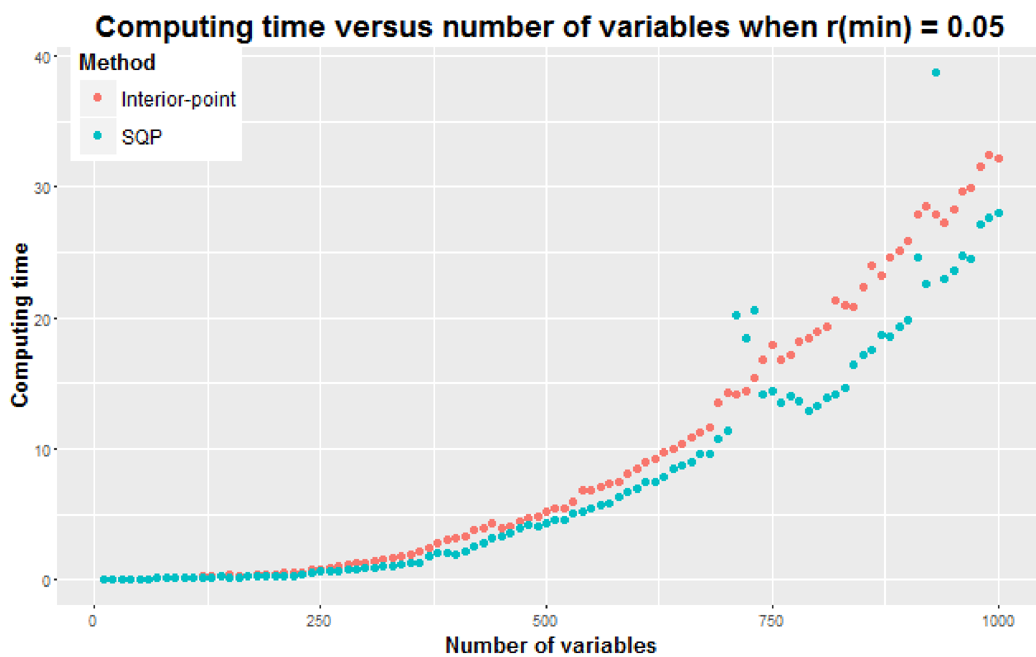


Figure 2: Computing time versus the number of available stocks: $r_{\min} = 0.05$

When $r_{\min} = 0.05$, although SQP method becomes slightly faster when n is larger than 750, the computing time between two methods are not very different. Note that there are some outliers under SQP method and the largest one is around $n = 900$. It proves the previous conclusion I made that SQP method has larger variance compared to interior-point method. It is clear that when $r_{\min}$ is relatively small, both methods work similarly well and the computing times are around the same. When n becomes even larger, the difference becomes slightly larger.

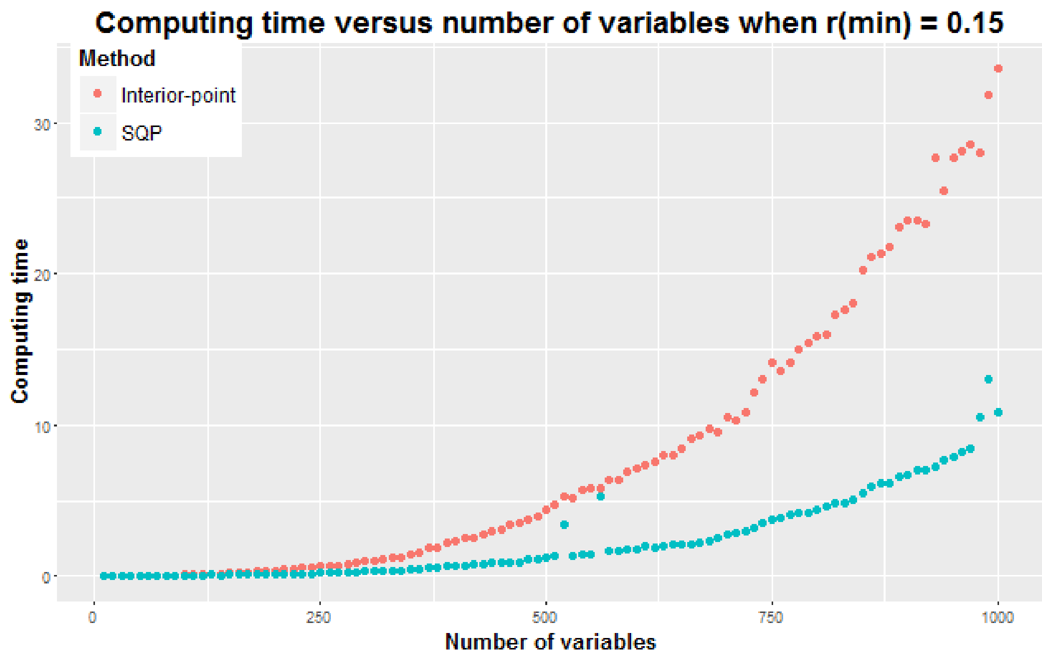


Figure 3: Computing time versus the number of available stocks: $r_{\min} = 0.15$

When $r_{\min} = 0.15$, the computing time between two methods starts to be well-distinguished, and SQP method runs much faster than interior-point method when n is greater than 500. The slope of the curve under interior-point method increases much faster than that under SQP method.

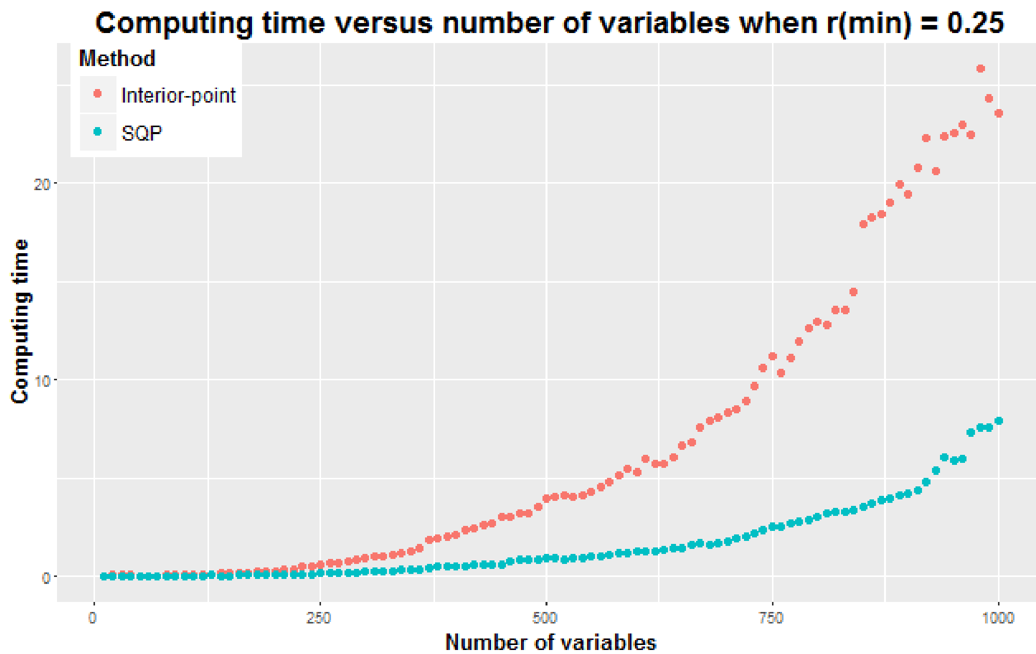


Figure 4: Computing time versus the number of available stocks: $r_{\min} = 0.25$

When $r_{\min} = 0.25$, the computing time between two methods is still very diversified, and SQP method runs much faster than interior-point method when n is greater than 500. Overall, the computing time of the two methods shows a slightly decreasing trend compared to the situation when $r_{\min} = 0.15$. Perhaps the more aggressive portfolio will become less diversified and decrease the iterative steps in both methods.

From the previous figures, we can mainly find out that when the minimum requirement of expected returns increases, both methods could solve the problem slightly quicker, but SQP method is more efficient in dealing with the optimization problem. On the other hand, the computing time under interior-point method is more consistent than the computing time under SQP method. Based on the tables and results shown above, I suppose when n becomes even larger to 10,000 and 100,000, the advantages of using SQP method will become more prominent. It could solve the problem much quicker than using interior-point method.

4 Conclusion, discussion and further research

In this paper, I mainly talk about some basic ideas of optimization problems and the importance of convexity. I use Markowitz portfolio management problem as an example to illustrate the difference between two widely-used solving methods: interior-point method and SQP method when the number of available stocks becomes large. In this example, I did a randomization test to examine the running time of two algorithms under different settings of the number of available stocks and minimum expected return. It turns out that as the number of available stocks increases, both algorithms will increase their computing time. However, if the minimum expected return is aggressive, SQP method become more efficient compared to interior-point method. The variance of computing time of interior-point method is less than the computing time of SQP method, which proves that interior-point method is more stable than SQP method.

In my research, I intended to do the comparison analysis when n is really large, such as 10,000 and 100,000. However, due to the limitation of computer hardware, the number of available stocks is limited to 1000, which is not very huge. If the computing power is very strong, I think the results when n is huge will become much more diversified. Another noted issue is the initialization of portfolio. It is possible that if there exists some different methods that have a better way to initialize the portfolio, they will save more computing time on both solving methods.

Speaking of further research areas, since I only did the comparative analysis in one time period, I think it is a good idea to simulate the stock trading in the real-world by doing a re-balancing of the portfolio over multiple time periods. If both algorithms need to run multiple times and evaluate the current portfolio, I think interior-point method with less variance will perform better than its performance in a single time period. Also, adding additional constraints, such as fixed and variable transaction costs, limits on the portion of a single asset or allowing short position are also some further approaches to this question.

References

[1] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[2] Harry Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952.

[3] William F Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442, 1964.

[4] John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The review of economics and statistics*, pages 13–37, 1965.

[5] B Blog, G Van Der Hoek, AHG Rinnooy Kan, and GT Timmer. The optimal selection of small portfolios. *Management Science*, 29(7):792–798, 1983.

[6] J David Jobson and Robert M Korkie. Putting markowitz theory to work. *The Journal of Portfolio Management*, 7(4):70–74, 1981.

[7] Narendra Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, pages 302–311. ACM, 1984.

[8] Robert R Meyer. Sufficient conditions for the convergence of monotonic mathematicalprogramming algorithms. *Journal of computer and system sciences*, 12(1):108–121, 1976.

[9] Nitin R Patel and Marti G Subrahmanyam. A simple algorithm for optimal portfolio selection with fixed transaction costs. *Management Science*, 28(3):303–314, 1982.

[10] Andre F Perold. Large-scale portfolio optimization. *Management science*, 30(10):1143–1160, 1984.

[11] Hideo Tanaka, Peijun Guo, and I Burhan Türksen. Portfolio selection based on fuzzy probabilities and possibility distributions. *Fuzzy sets and systems*, 111(3):387–397, 2000.

[12] Olivier Ledoit and Michael Wolf. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, 10(5):603–621, 2003.

[13] Lorenzo Garlappi, Raman Uppal, and Tan Wang. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies*, 20(1):41–81, 2007.

[14] Jun Liu. Portfolio selection in stochastic environments. *Review of Financial Studies*, 20(1):1–39, 2007.

[15] Campbell R Harvey, John C Liechty, Merrill W Liechty, and Peter Müller. Portfolio selection with higher moments. *Quantitative Finance*, 10(5):469–485, 2010.

[16] Robert B Wilson. *A simplicial algorithm for concave programming*. Graduate School of Business Administration, George F. Baker Foundation, Harvard University, 1963.